Quick Guide to EduG 6.0

Download from website: http://www.irdp.ch/edumetrie/englishprogram.htm

After downloading complete the following instructions to be compatible with security software:

This bug is a side effect of an ''excessive'' protection of Windows . EduG uses an external module (APLgrid.dll) which must be **Registered** by Windows. The registration is automatically done by the install procedure and should be unchanged. Unfortunately I believe that some '' Cure System Cleaners'' remove Registry of ''APLgrid.dll" because it is potentially dangerous!

Since Windows Vista this Registration is protected and can only be done by the Administrator of the computer. If you are running Windows Vista , try the following actions (**Attention**: You must be running your computer as **'' Administrator ''**)

1. open Windows Control Panel
2. select User Accounts
3. click "Turn User Account Control on or off"
4. clear the check box labeled "Use User Account Control (UAC) to help protect your computer"
5. click OK
6. restart computer for this change to take affect

Then EduG will automatically register the module ''APLgrid.dll''\

If you are running Windows 7, the actions should be the same: you have to deactivate the User Account Control (UAC). The Windows7 panels are different of those of Windows Vista (I have only the french version of Windows 7 ... I give you an approximative translation in English):

1. open Windows Control Panel
2. select Protection of User Accounts

3. select Modify your User Accounts

4. select Modify Control Parameters of User Account

5. **move cursor down**

6. Click **OK,** then Close Control Panel

7. Restart computer for this change to take effect

Then EduG will automatically register the module ''APLgrid.dll''

USER MANUAL: http://www.irdp.ch/edumetrie/documents/EduGUserGuide.pdf

Data Constraints

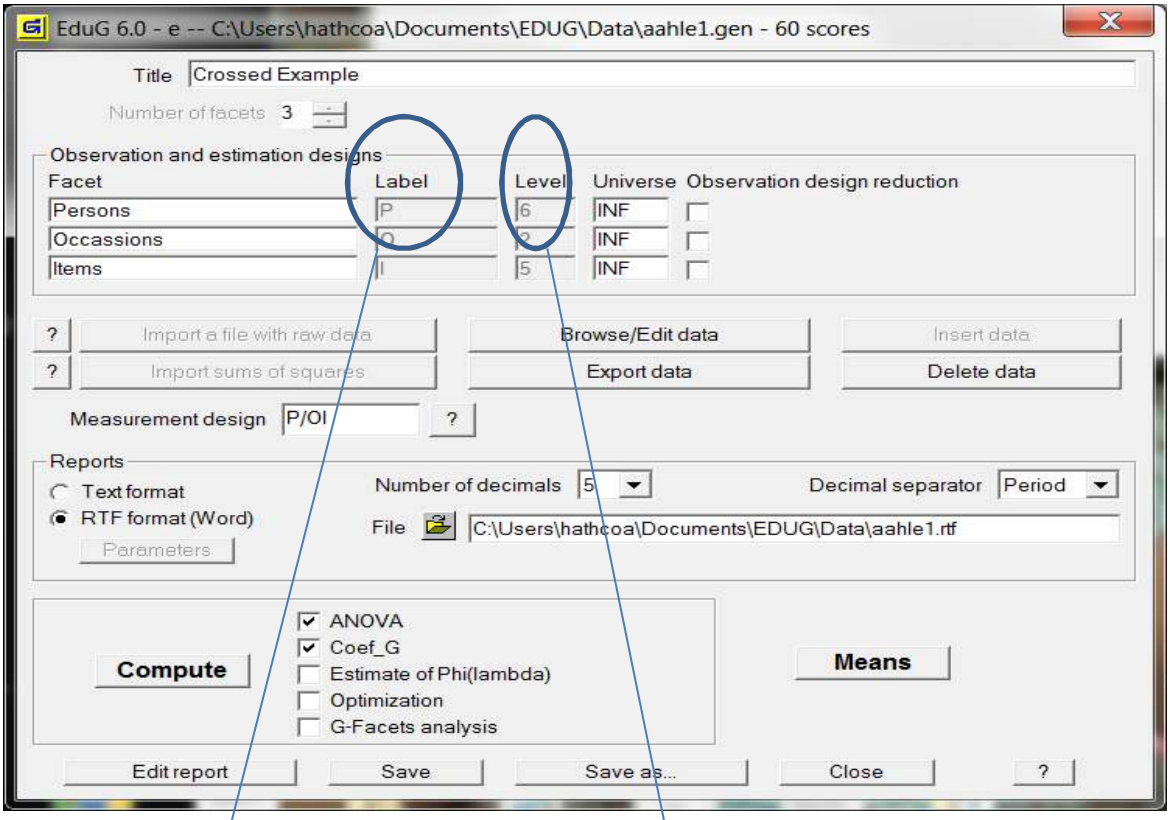
1. If measures are made using several items then all measurements should be on the same scale (e.g. 1-5).
    - If this is not the case then you may force them to be on the same scale or you may choose to report proportions (1/5, 2/5/ 3/5…etc).
2. EduG 6.0 can only handle balanced data. In other words if students are nested within classes then we must have equal students in each class. If raters are nested within occasions then the number of raters assigned to each occasion must be equal.
    - Consider data imputation methods
    - Force design to be balanced (randomly select to obtain balanced design)
    - Estimate sums of squares using SAS or SPSS and then use sums of squares to estimate variance components in EduG 6.0.
3. EduG 6.0 can handle up to 8 facets, which includes a differentiation facet (what we wish to distinguish in a measurement procedure).

Determining Observation and Estimation Design

This is perhaps the most important aspect of using the EduG 6.0 program. The following steps should be taken when conducting a G-study:

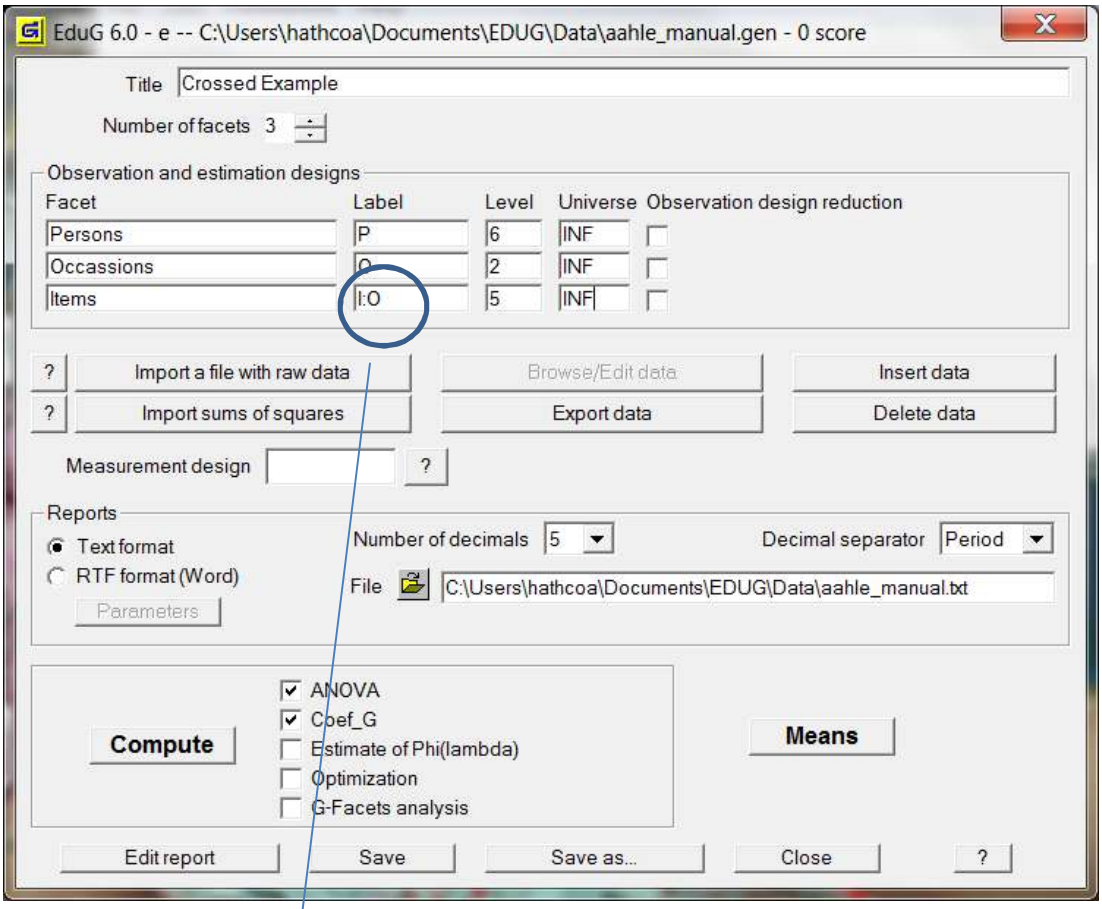
1. The first thing to consider is identifying the object of differentiation or the differentiation facet. In other words, what do we wish to differentiate? Are we concerned about examining differences among students, items, methods, etc?
2. Other facets will contribute to error in our efforts to differentiate our object of measurement/differentiation.
3. Now we must first identify each facet (including the differentiation facet) and determine the number of levels for each facet. If there are two raters in each occasion then raters will have two levels. If 30 people are tested on two occasions then a “person” facet will have 30 levels. It is important in most situations to first name the facet which changes least rapidly in your data file. Then identify the face that changes the next least rapidly and so on. This is further explained below under “Importing Data”.
4. After we identify the facet we must then provide a label. The label is also where we identify whether a facet is nested within a second facet. Remember if facet A has two or more levels associated with facet B then A is nested within B. If 20 people are in a class and there are 5 classes then persons are nested within classes only if we have different people in each classroom. It is conceivable for classrooms to be nested within students, though this study is typically rare. For example let's assume that we have 8 children that are observed within 3 classrooms. If these children are attending different courses then classrooms are nested within children. If in a different study we have 2 groups of raters and let’s assume that there are three raters within each group. One group of raters is assigned to classroom 1 and the second group is assigned to classroom 2. In this case, raters are nested within classrooms. Raters would be crossed with classrooms if every rater went to each classroom.

So let's assume that we wish to design a study that is fully crossed. For this study we have 6 students observed by a group of raters across 2 occasions. In this situation, each rater (i.e. 2 total raters) is assigned to each occasion. Consequently this is fully crossed given that each rater observes each student for every occasion. The data for this situation is actually presented below under "importing data". However, for now we will examine how this design is specified in the EduG 6.0 program.



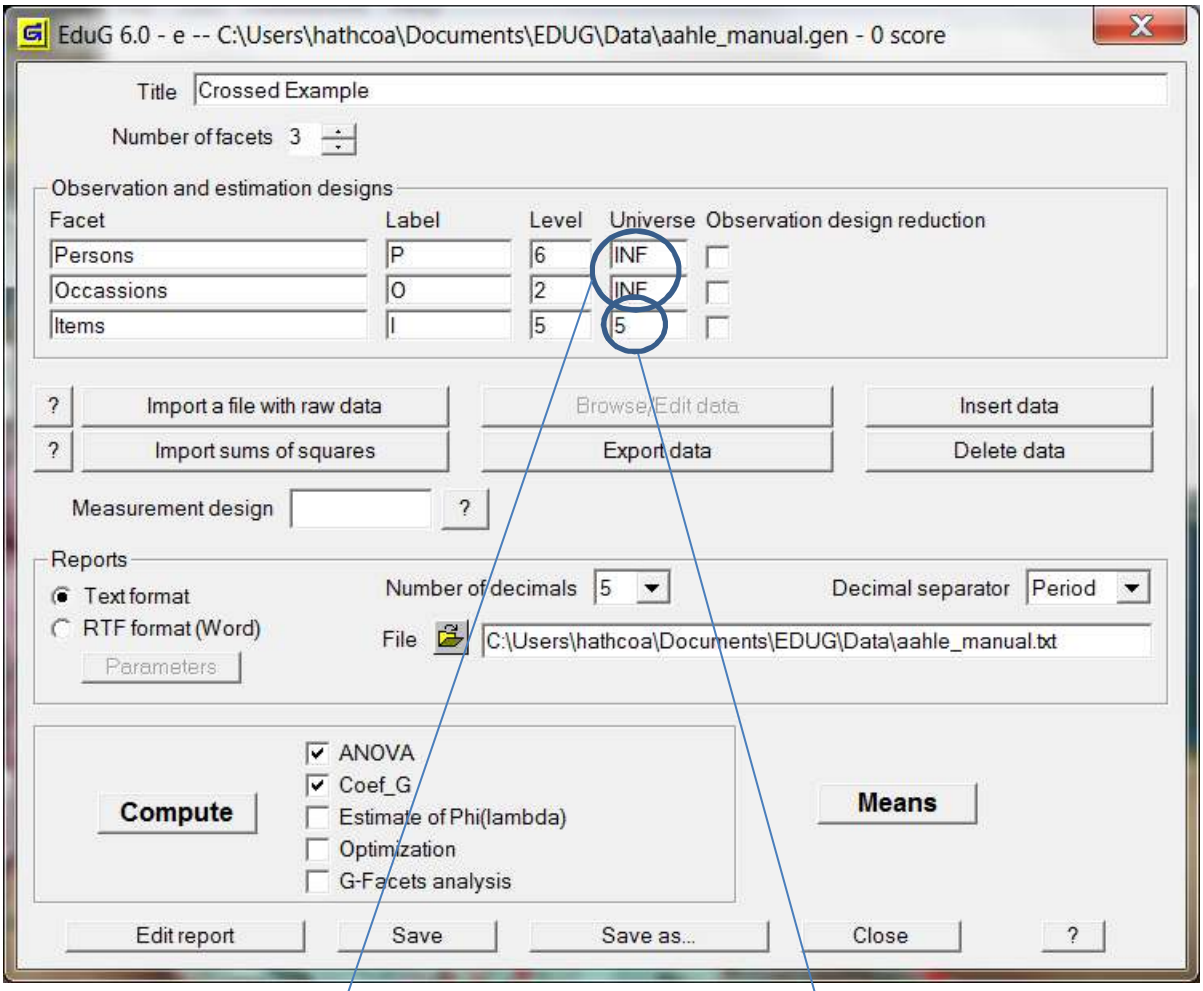
Simply use a capital letter to signify each Facet.

There are 6 persons observed on two occasions. On each occasion each person responded to the same 5 items.



If we had a different design (on each occasion individuals responded to different items then items would be nested in occasions).

I:O indicates that items are nested within occasions. This is shown only for illustration purposes.



This indicates that the universe is infinite for these facets. In other words, we are willing to treat both persons and occasions as interchangeable. Other raters and occasions of the same sample size would work equally well.

In order to illustrate a "fixed" facet I have inserted a 5 here. This indicates that we are only interested in these 5 items. This limits our generalization claims to these specific items, but this choice would eliminate items as a source of error.

Though in the previous example I stated that items are fixed we are actually interested in inferences pertaining to any items of the same characteristics. In other words, for this example we will treat items as interchangeable or random.

The next step is to import data. This can be done in several ways (import raw data file, insert data manually, or insert sums of squares). I will briefly review how to import data using a raw data file, though I generally find the "insert data" command to be more useful.

Importing Data

1. Important to convert file to ASCII format before importing.
2. To be successful care must be taken that facets declared in EduG 6.0 conforms to structure of the file.
    - The first facet declared in EduG 6.0 should be one whose levels change least rapidly when scanned from left to right.
    - The last facet to be declared would be the facet whose levels change most rapidly when scanning from left to right.

Let's assume that we have 6 persons measured on 2 occasions. At each occasion each person responded to the same five items. This may have a format that is similar to the data given below. For this data, persons are changing least rapidly, followed by occasion and then item. Consequently in the EduG 6.0 program it is important to first label persons, then occasions, and then items.

This file was pulled from excel. To actually import the file however, I've noticed that you must delete the first row indicating variable names and the person, occasion, and item column. The program will then only read the "score" column (of course the name score is deleted). So before saving you should delete everything except the actual scores. If you have appropriately labeled the facets in the EduG 6.0 program (putting in person first, then occasion, then item) then the program will accurately read a single column of data.

You can also insert the data manually after defining your facets. I have often found it easier to simply click "insert data" and then paste the values from an excel file into the EduG 6.0 program. This is also a good way to examine whether you are thinking about your design appropriately.

Person	Occasion	Item	Score
1	1	1	5
1	1	2	5
1	1	3	4
1	1	4	4
1	1	5	3
1	2	1	5
1	2	2	5
1	2	3	4
1	2	4	4
1	2	5	4
2	1	1	3
2	1	2	4
2	1	3	2
2	1	4	3
2	1	5	2
2	2	1	5
2	2	2	5
2	2	3	2
2	2	4	2
2	2	5	4
3	1	1	1
3	1	2	4
3	1	3	5
3	1	4	2
3	1	5	4
3	2	1	2
3	2	2	3
3	2	3	5
3	2	4	2
3	2	5	3
4	1	1	3
4	1	2	3
4	1	3	2
4	1	4	2
4	1	5	4
4	2	1	4
4	2	2	4
4	2	3	2
4	2	4	2
4	2	5	2
5	1	1	1
5	1	2	4
5	1	3	4
5	1	4	2

Delete this row before saving as tab delimited file

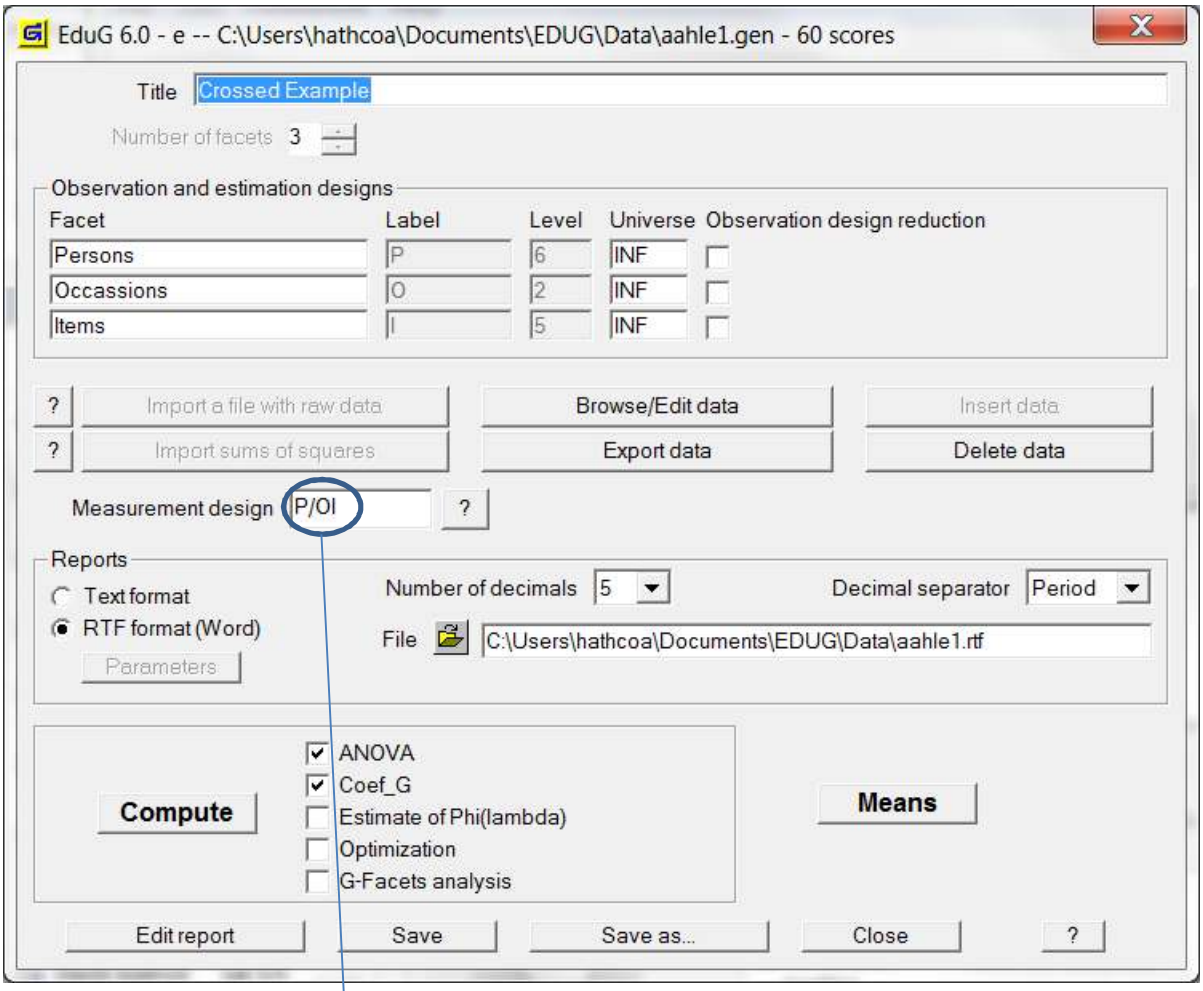
Delete these three columns before saving as tab delimited file

5	1	5	2
5	2	1	2
5	2	2	4
5	2	3	4
5	2	4	3
5	2	5	5
6	1	1	2
6	1	2	2
6	1	3	2
6	1	4	1
6	1	5	2
6	2	1	1
6	2	2	3
6	2	3	1
6	2	4	1
6	2	5	2

Once the data is imported then the file is automatically saved. After saving we can no longer change the observation design (i.e. FACET LABELS OR LEVELS). We can however, still change whether a facet is declared as random or finite.

Measurement Design

This section indicates the specification of a measurement design.



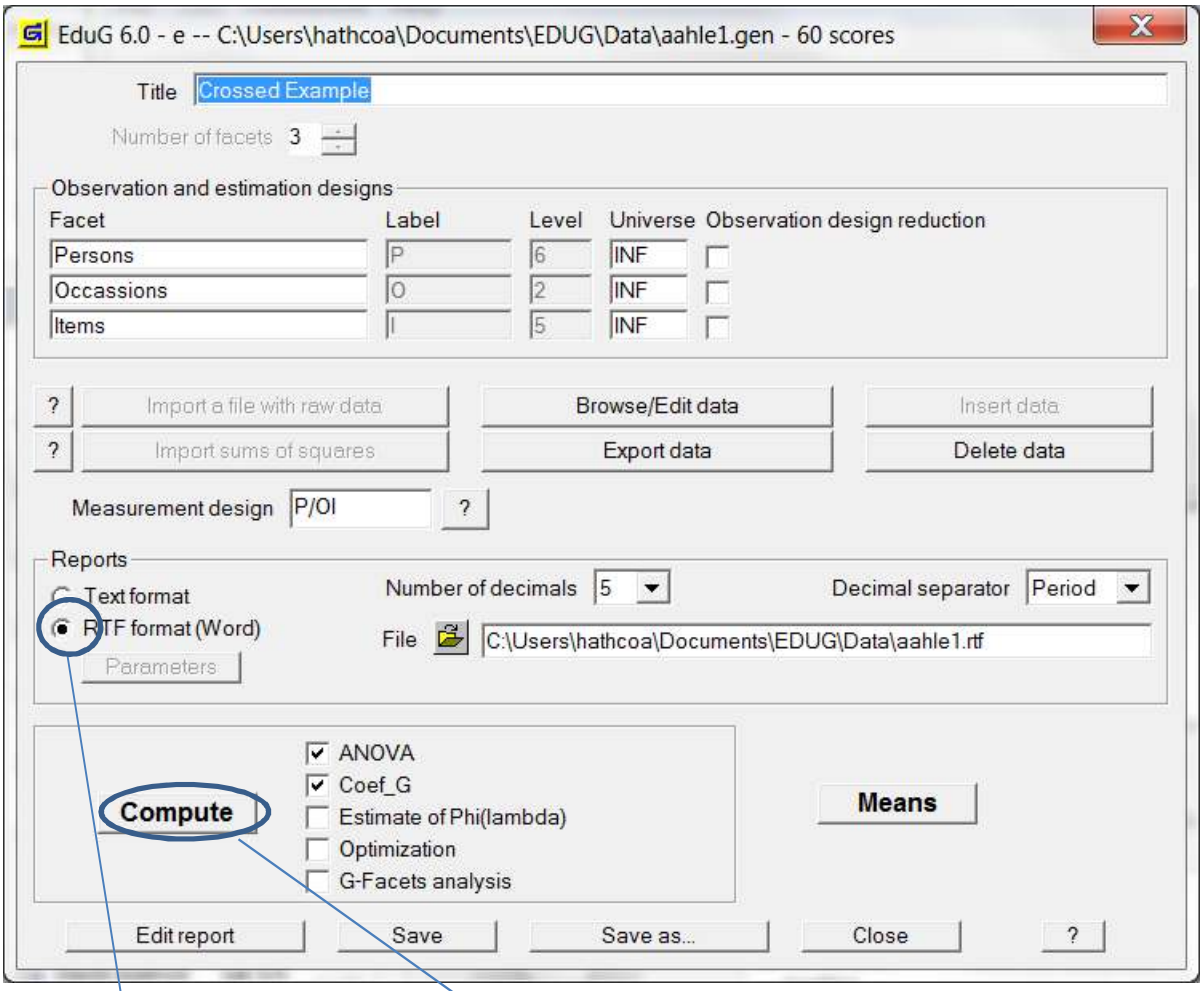
A / is used to indicate differentiation facets (object of measurement) from sources of error. Differentiation facets are placed to the left of / while sources of error are placed to the right.

Some points that are worth noting:

1. Do not indicate nesting in the measurement design. We have already indicated what facets are nested so there is no need to replicate that information here.
2. If the object of differentiation is nested within a second facet then you must include both of these facets on the left hand side of "/" . So for example, if persons were nested in occassions then our measurement design would be PO/I.
3. The inverse of this rule is not true. So in other words, if we have a facet that is nested in the differentiation facet (object of measurement) then there is no need to include this information to the right of the "/". So for example, if items are nested in persons then items (error) are nested within a differentiation facet (persons). Consequently our measurement design would remain: P/OI.

Interpretation of Output

The following section will briefly illustrate how to compute and interpret output for the person x occasion x item design. There are numerous other examples within the EduG 6.0 manual and that can be found in Cardinet, Johnson, & Pini, 2010.



Click this option if you wish to have output/reports in word.

Once you have imported the data and have specified a measurement design you may then click "compute" to generate a report.

You will actually see more output than what is displayed below. I will however, provide you with a quick description of the output that is most relevant.

Analysis of Variance

Source	SS	df	MS	Components				
				Random	Mixed	Corrected	%	SE
P	34.73333	5	6.94667	0.53333	0.53333	0.53333	29.8	0.37615
O	1.06667	1	1.06667	0.02000	0.02000	0.02000	1.1	0.03260
I	14.10000	4	3.52500	0.18500	0.18500	0.18500	10.3	0.17518
PO	3.33333	5	0.66667	0.02167	0.02167	0.02167	1.2	0.07882
PI	30.10000	20	1.50500	0.47333	0.47333	0.47333	26.4	0.24200
OI	1.43333	4	0.35833	-0.03333	-0.03333	-0.03333	0.0	0.04445
POI	11.16667	20	0.55833	0.55833	0.55833	0.55833	31.2	0.16834
Total	95.93333	59					100%	

This lists the variance components for our estimation design. It is important to understand what these components reflect for later interpretation. So for example, P is differentiation facet and reflects mean differences of each person across each occasion and item. PO indicates the extent to which person scores change across each occasion. A high PO would suggest that which persons were ranked higher changed across occassions.

This indicates the relative proportion of variation attributed to each facet or combination of facets. Notice that PI and POI are relatively large.

Notice that there are three columns: Random, Mixed, and Corrected. Since all of our facets are random we are using a random model. However, if some facets were fixed we would use the mixed model. Only use the corrected model if the differentiation facet is fixed.

G Study Table

Source of variance	Differ-entiation variance	Source of variance	Relative error variance	% relative	Absolute error variance	% absolute
P	0.53333		.....		.....	
	.....	O	.....		0.01000	4.8
	.....	I	.....		0.03700	17.8
	.....	PO	0.01083	6.7	0.01083	5.2
	.....	PI	0.09467	58.7	0.09467	45.4
	.....	OI	.....		(0.00000)	0.0
	.....	POI	0.05583	34.6	0.05583	26.8
Sum of variances	0.53333		0.16133	100%	0.20833	100%
Standard deviation	0.73030		Relative SE:	0.40166	Absolute SE:	0.45644
Coef_G relative	0.77					
Coef_G absolute	0.72					

This is the differentiation variance (universe score variance). If this is very small, then a measurement procedure may have difficulty detecting differences (i.e. there are little differences to detect)

This is important if we are interested in absolute decisions (i.e. placing students on a scale). Notice that all sources of error are included in this decision whereas only some sources of error affect relative rankings.

This is important when making relative decisions (which students are higher or lower than others). This is only affected by interactions of the differentiation facet with other facets (PO, PI, etc)

Source of variance	Differ-entiation variance	Source of variance	Relative error variance	% relative	Absolute error variance	% absolute
P	0.53333		.....		.....	
	.....	O	.....		0.01000	4.8
	.....	I	.....		0.03700	17.8
	.....	PO	0.01083	6.7	0.01083	5.2
	.....	PI	0.09467	58.7	0.09467	45.4
	.....	OI	.....		(0.00000)	0.0
	.....	POI	0.05583	34.6	0.05583	26.8
Sum of variances	0.53333		0.16133	100%	0.20833	100%
Standard deviation	0.73030		Relative SE:	0.40166	Absolute SE:	0.45644
Coef_G relative	0.77					
Coef_G absolute	0.72					

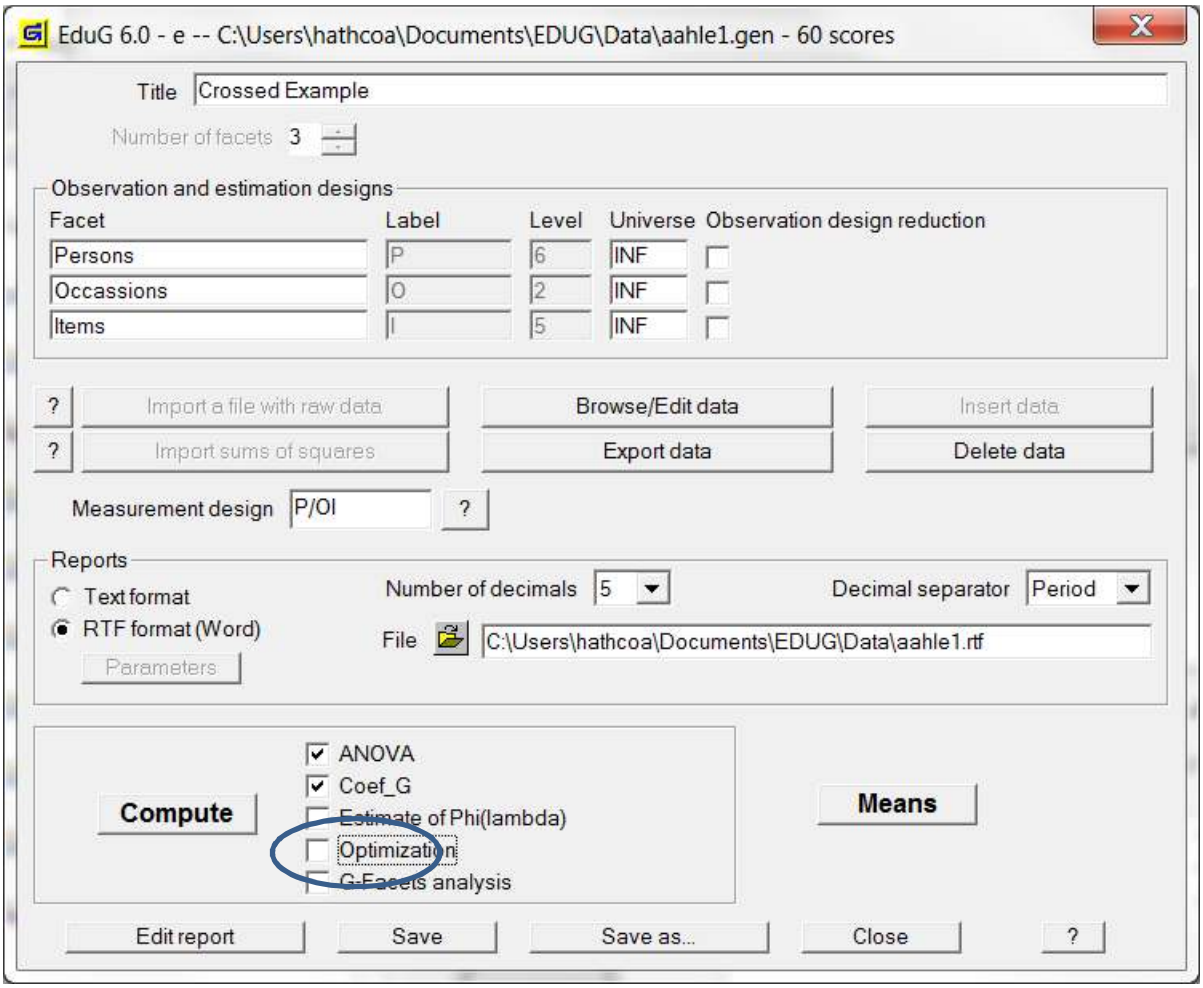
This indicates the two reliability-like coefficients associated with making both relative and absolute decisions. In most cases the absolute will be lower than the relative coefficient.

This is an important piece of information. It indicates how much scores are expected to vary if the study were replicated by taking a random sample of 5 items and 2 occasions. In other words, on average each person's score tends to deviate .40 points from their universe score (what would be expected across all items and occasions.

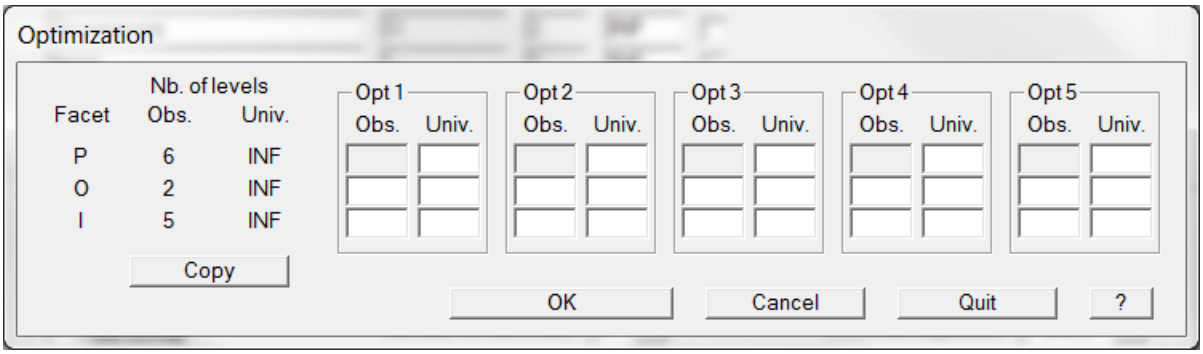
We can also use this to create confidence intervals around individual scores. For 95% CI simply: $Score \pm SE(1.96)$

D-Study (Optimization Study)

A Decision study (i.e. D-study) allows us to use the information from a generalizability study in order to examine how changes in sampling may influence the results. Below is a brief illustration depicting how to use the optimization feature of the EduG 6.0 program.



After checking the circled box you will be presented with the following screen. We can modify the options in several ways in order to examine the effect of increasing the number of items or occasions on the estimated generalizability coefficient. In this situation, since a large portion of the error variance resides within PI we will examine the effect of changing items. It should be noted however, that we can try different combinations of items and occasions in order to optimize our measurement procedure.



Below I have kept everything constant except for the number of items.

Optimization

	G-study		Option 1		Option 2		Option 3		Option 4		Option 5	
	Lev.	Univ.	Lev.	Univ.	Lev.	Univ.	Lev.	Univ.	Lev.	Univ.	Lev.	Univ.
P	6	INF	6	INF	6	INF	6	INF	6	INF	6	INF
O	2	INF	2	INF	2	INF	2	INF	2	INF	2	INF
I	5	INF	6	INF	7	INF	8	INF	9	INF	10	INF
Observ.		60		72		84		96		108		120
Coef_G rel.		0.76775		0.79652		0.81841		0.83565		0.84956		0.86103
rounded		0.77		0.80		0.82		0.84		0.85		0.86
Coef_G abs.		0.71910		0.75073		0.77509		0.79441		0.81013		0.82315
rounded		0.72		0.75		0.78		0.79		0.81		0.82
Rel. Err. Var.		0.16133		0.13625		0.11833		0.10490		0.09444		0.08608
Rel. Std. Err. of M.		0.40166		0.36912		0.34400		0.32388		0.30732		0.29340
Abs. Err. Var.		0.20833		0.17708		0.15476		0.13802		0.12500		0.11458
Abs. Std. Err. of M.		0.45644		0.42081		0.39340		0.37151		0.35355		0.33850

Note how the estimated G_coefficient changes as we increase the number of items. This information allows one to conclude that increasing the number of items to 7 would provide a generalizability coefficient approximately .82. It is up to you however, as a researcher to decide what optimization procedure is appropriate given practical constraints. In this situation, we may be satisfied with the original G_coefficient and decide that increasing the number of items for this improvement is not worth additional resources.

Variance Attribution Diagrams

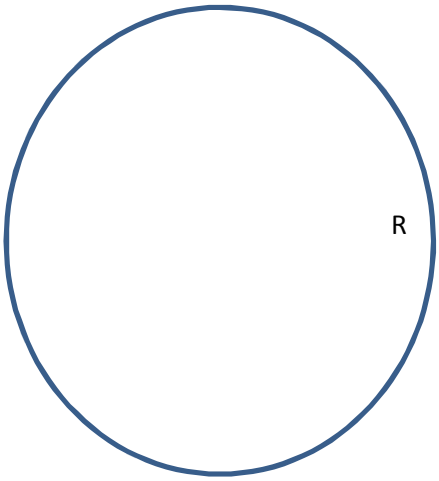
Variance attribution diagrams are extremely beneficial before conducting a G-study. These diagrams allow you to determine which sources of error may be confounded. These diagrams are also beneficial in that they allow you to determine which sources of error contribute to both relative and absolute decisions. Unfortunately, I was unable to find a reference that described how to construct these diagrams in a way that is easy to understand. The book provided by Cardinet and colleagues (2010), as well as the Shavelson and Webb (1991) introduction to G-theory, are good places to start. They do discuss these in slightly different terms however, though I believe that the presentation by Cardinet et al. (2010) is easier to follow than the Shavelson and Webb (1991) text.

These diagrams can become increasingly complicated, particularly for designs with numerous facets. It is also important to recognize that specifying a facet to be fixed can drastically alter the sources of error that can be estimated in a G-study. As a general rule of thumb, if a facet is fixed then it will not contribute to error. This makes sense given that this specification indicates that we have observed every possible level (the universe of admissible observations is contained in our measurement procedure) thus it will not contribute to error when we attempt to make generalized inferences. I would suggest that you consult both of these texts to get a better understanding of fixed facets, and the construction of Venn diagrams in this situation. I will however, provide an overview of two designs; one of which is completely crossed and the other has a facet that is nested within a second facet. I hope that the presentation provided below, when coupled with the information provided by these other authors, will allow you to construct Variance Attribution Diagrams when thinking about the best design for your study.

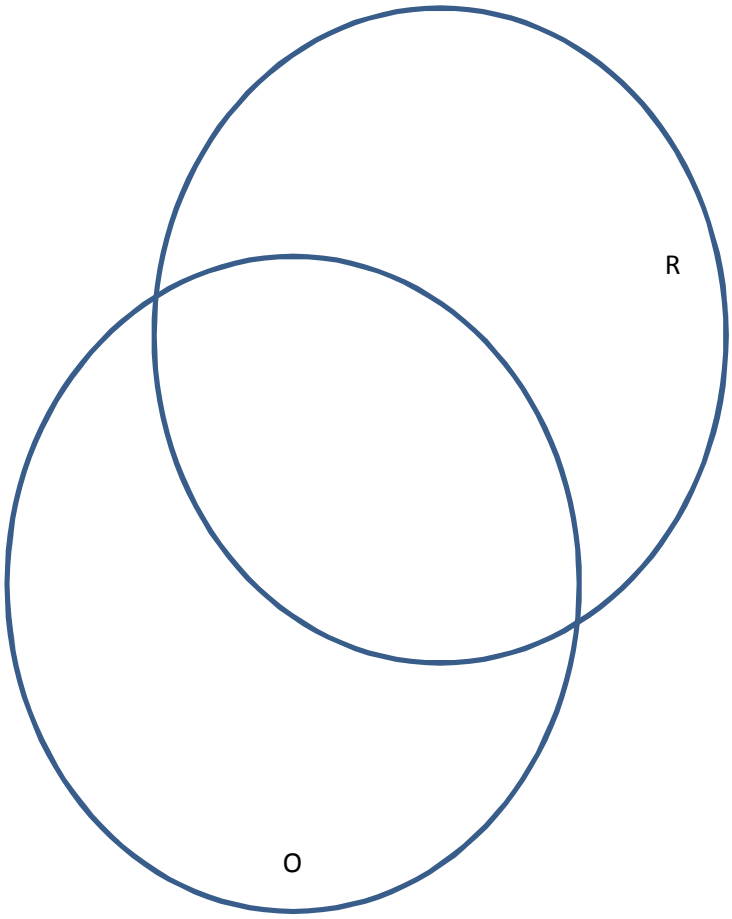
Variance Attribution Diagram – Crossed Design

To remain consistent, I will present a variance attribution diagram that is structured around the previous example. In this example we have 6 students that are observed on 2 occasions. All raters observe each student on every occasion. Thus we have a person x occasion x rater design.

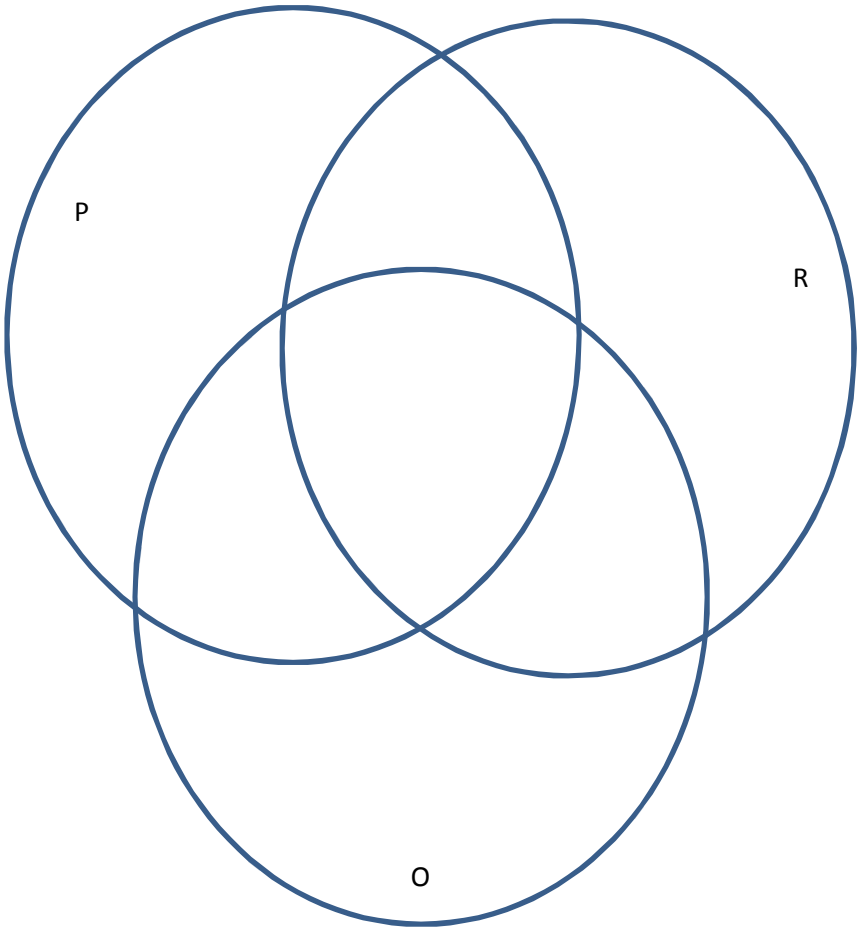
Step 1: Make a circle to signify raters (the choice in this situation is rather arbitrary).



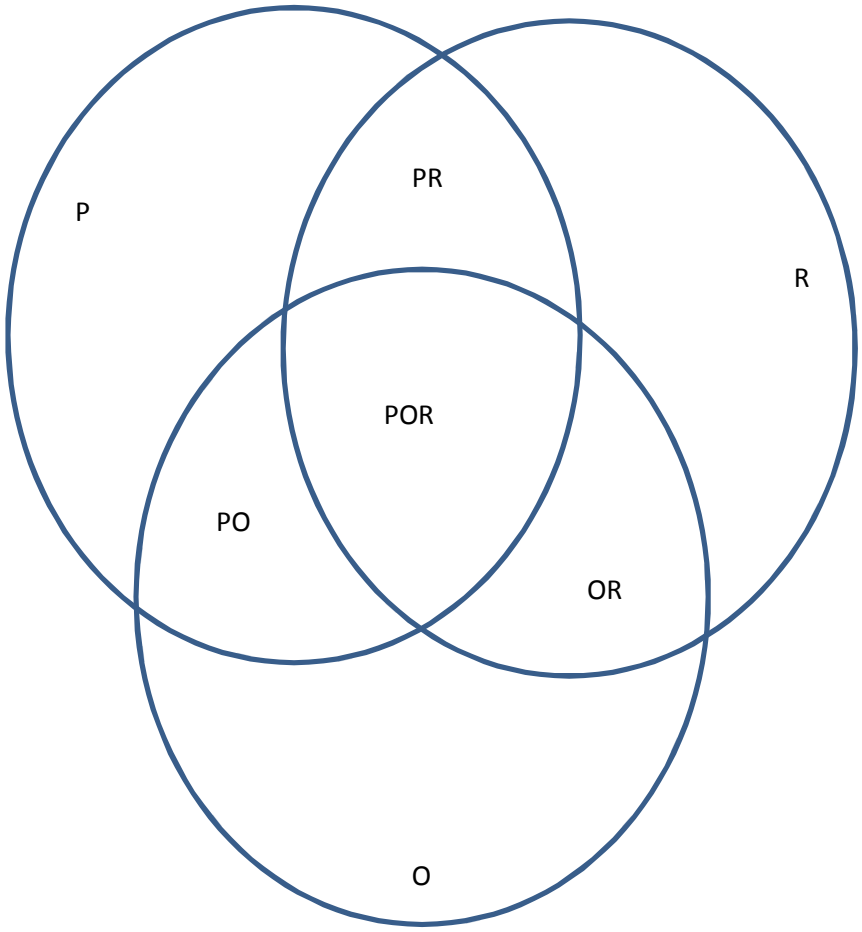
Step 2: Draw a circle to represent occasions. Since occasions are crossed with raters the circles must intersect.



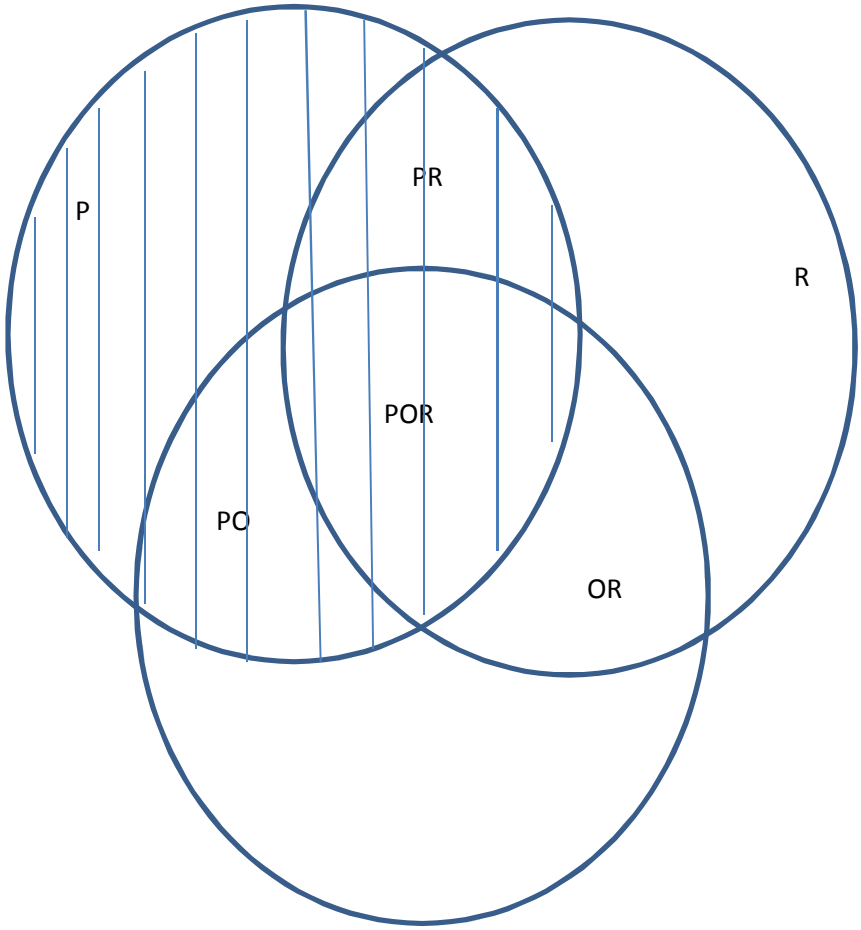
Step 3: Now do the same thing for persons. Since persons are crossed with both raters and occasion then these circles should also intersect.



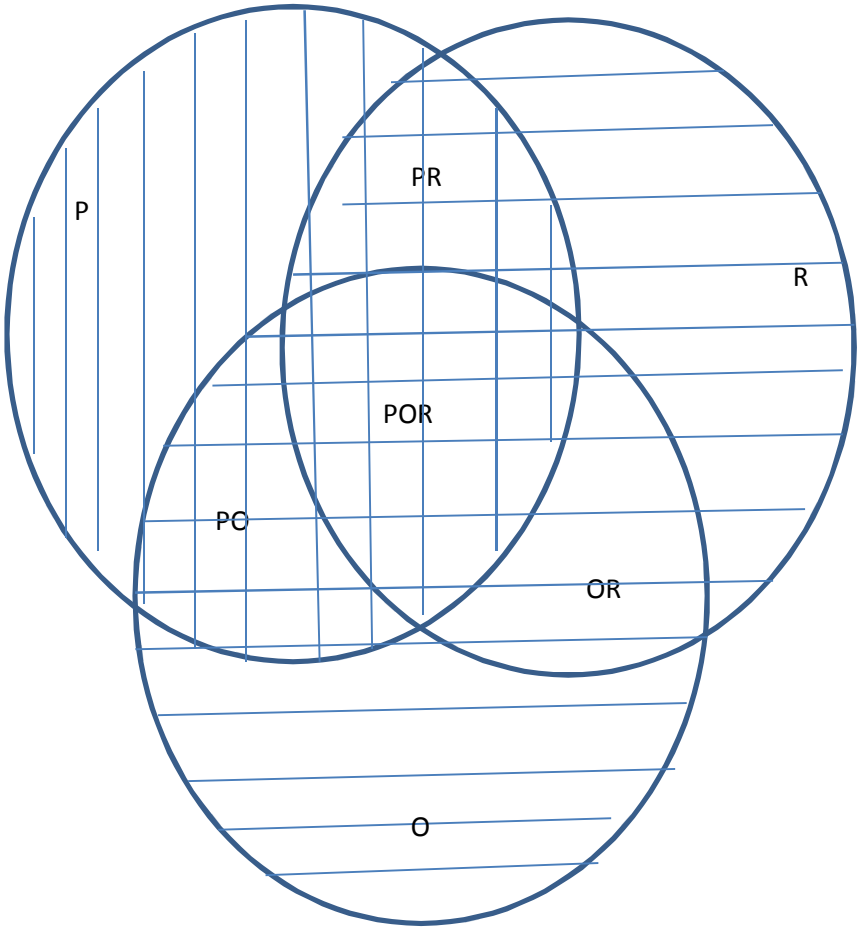
Step 4. Now we should label each intersection.



Step 5: Identify the differentiation facet / object of measurement. In this case we wish to differentiate people, though note that we could easily use this design to differentiate raters or occasions. Create vertical lines through the differentiation facet.



Step 6: Now create horizontal lines through the facets of differentiation / facets that potentially contribute to error in our measurement procedure.

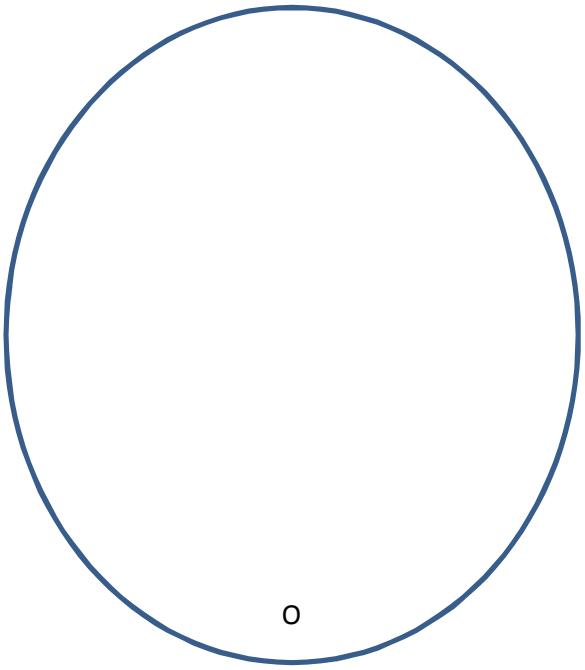


KEY – Hatched areas signify error pertaining to relative decisions. Error for relative decisions is therefore any point within the object of differentiation at which an intersection exists with other facets. Error for absolute decisions includes both the hatched lines and the horiziontal lines (i.e. R, OR, and O).

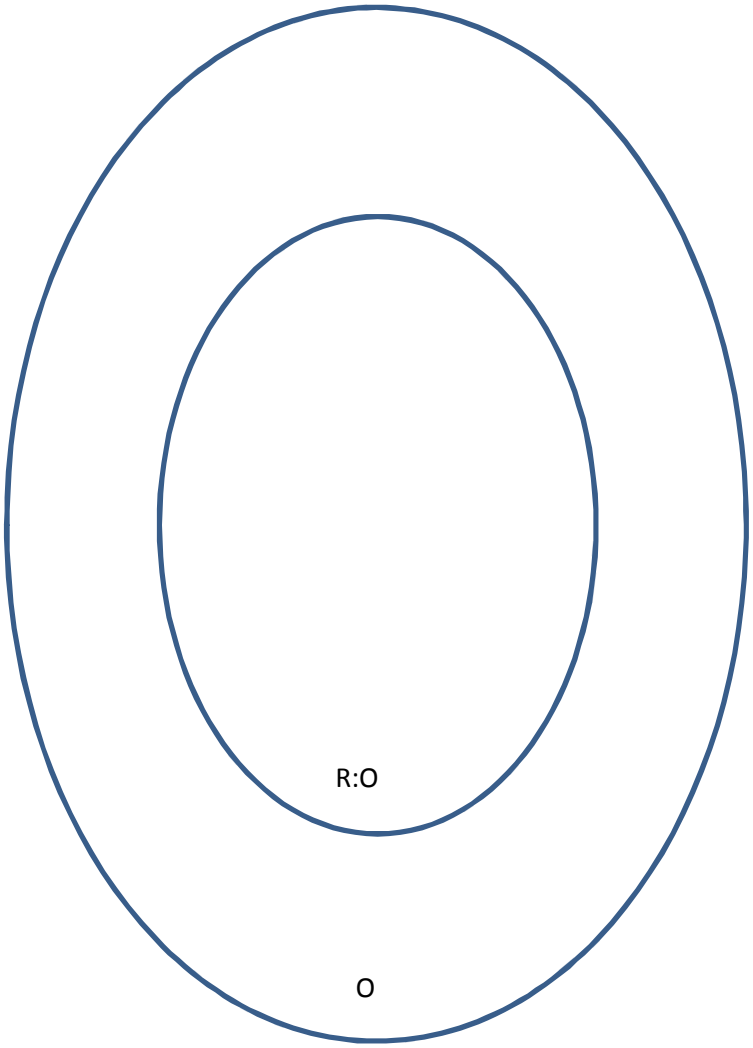
Variance Attribution Diagram – Nested Design

For this design we will modify the previous example, so that we may compare it with the fully crossed design. Let's assume that we assigned 2 raters to each occasion. There are different raters however, for each occasion. In this case we have 6 individuals that are observed on 2 occasions. The main distinction is that now we have 4 raters (2 assigned to occasion 1 and 2 assigned to occasion 2). This indicates that raters are nested within occassions.

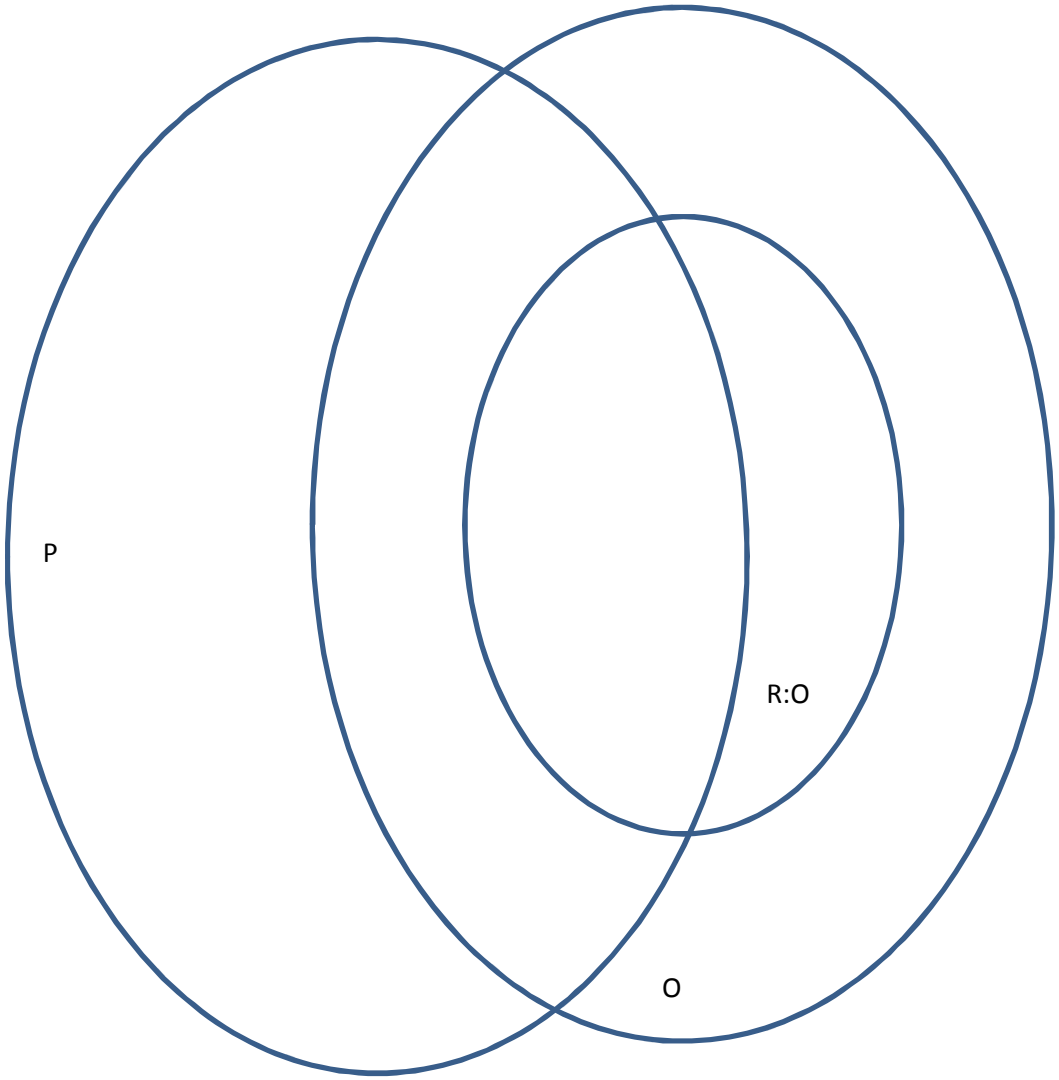
Step 1: Let's start with the occasion circle.



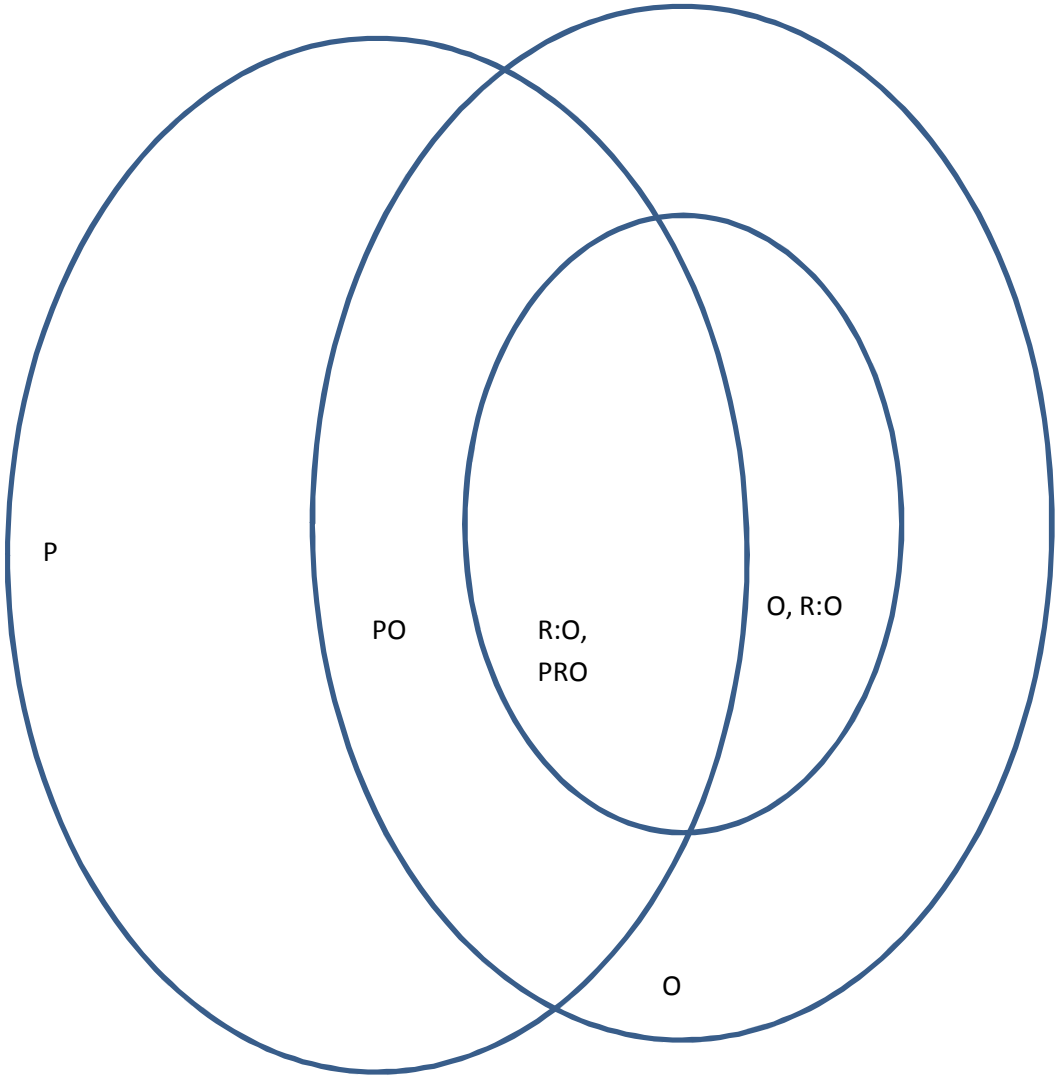
Step 2: Now we have raters nested within occasions. In order to indicate nesting we place the nested facet inside whatever it happens to be nested within. Since raters are nested within occasions we will therefore place the circle for raters inside the circle for occasions.



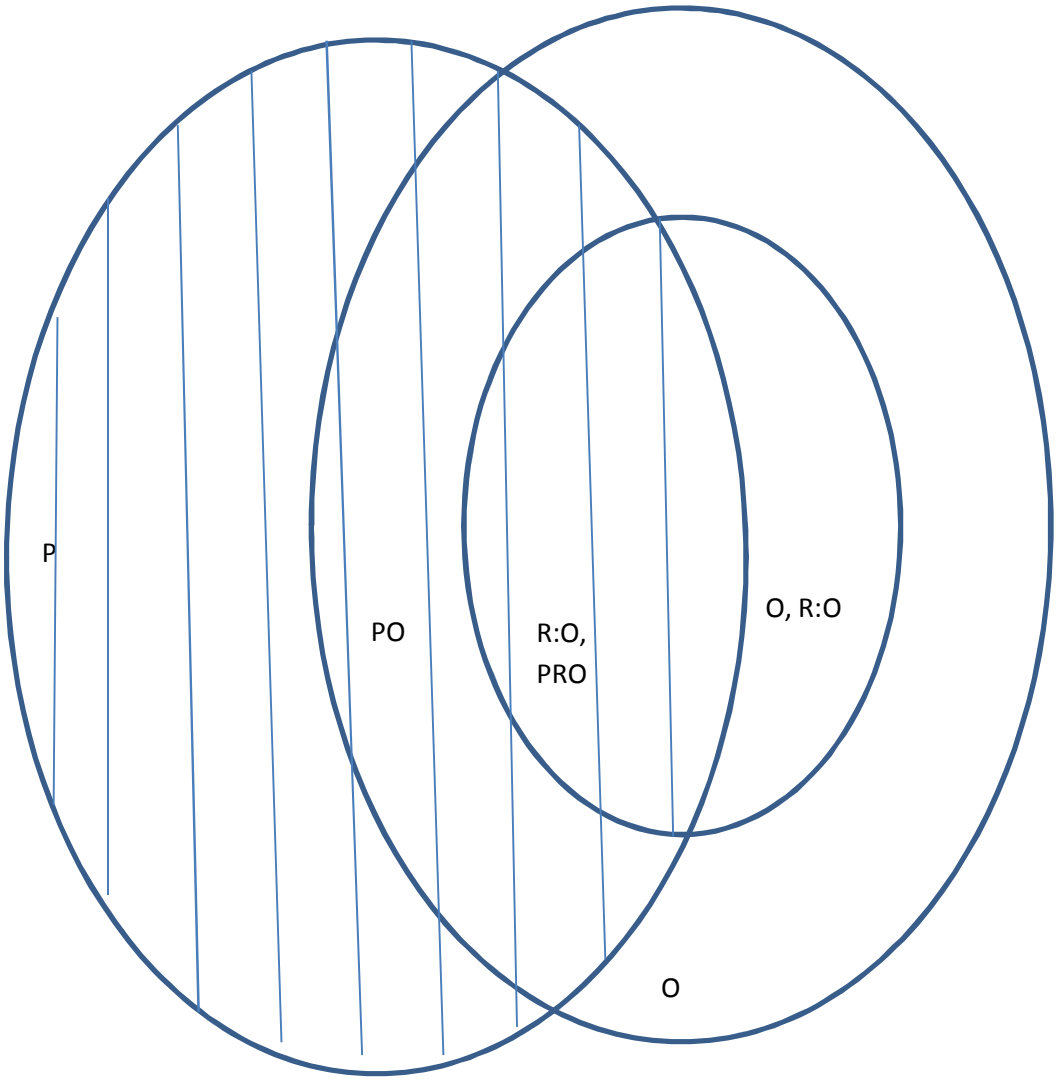
Step 3: We now have to indicate a circle for person variance. Since all people are examined across each occasion then this circle should intersect occasions. It is also important to note that all people are rated by the same raters. Consequently this circle should intersect with both the O and the R:O circles.



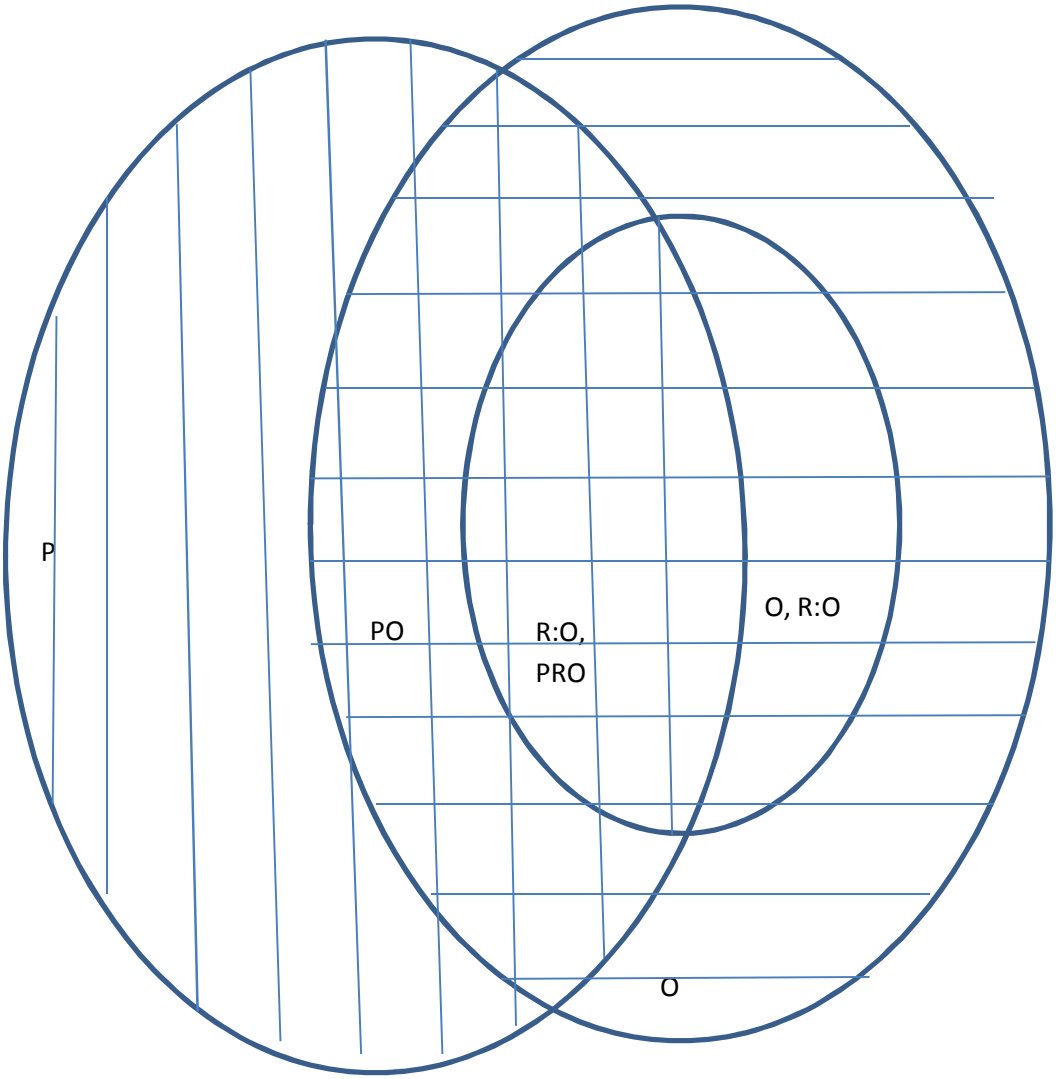
Step 4: Let us now label each intersection. It is important to note that we must still label and intersection among O and R:O. In other words, this intersection will have both O: and R:O and consequently R:O (effect of raters within each occasion) is confounded with an occasion effect. In other words we cannot estimate the effect of raters within each occasion separately from an occasion effect. It is also important to indicate that there is a 3 way interaction (PRO) that potentially exists at the intersection of PRO. Though we can estimate an effect of R:O it cannot be estimated separately from a three way interaction. Consequently, we will have output for R:O but this is confounded with a three-way interaction.



Step 5: Now we will create vertical lines within the circle that represents our facet of differentiation or the object of measurement. In this case we will again place vertical lines within the P circle since we wish to differentiate people.



Step 6: Now create horizontal lines across the sources of error that we wish to make generalizations. In this case we will draw horizontal lines across both R:O and the O circles.



KEY – Once again the hatched areas indicate sources of error for relative decisions. Hatched areas + areas with a horizontal line indicate sources of error for absolute decisions.

Cross or Nested Designs?

1. As you can tell the fully crossed design is in many ways optimal to a nested design. Nesting leads to confounding sources of error (i.e. O, and R:O).
2. Fully crossed designs therefore allow us to examine more sources of error than nested designs.
3. There are practical advantages to nesting however. Some facets may be naturally nested within others (i.e. students nested within classrooms) whereas at other times we can force nesting (assign different raters to each occasion).
4. There may be other advantages to nesting, such as better control over carry-over effects. For example, what would be the impact of subjecting the same students to the same items multiple times? Might a practice effect be at work? Carry-over effects can partially be controlled through nesting.
5. General advice – use fully crossed designs when feasible.